

Ari Vítor S. Lazzarotto

Capao do Leao, RS, Brazil | +55 53 99932-3611

arivitordasilva@gmail.com | linkedin.com/in/ar1vit0r | github.com/ar1vit0r

Summary

AI Engineer with hands-on experience building autonomous agents and Generative AI solutions using LangChain, LangGraph, and LLMs. Skilled in prompt engineering, LLMOps practices, and cloud-based AI architectures (AWS, Azure). Proven ability to design, deploy, and monitor intelligent agent systems with production-grade observability and security.

Experience

Data Scientist (Remote)

Paipe | Sep 2024 - Present

- Lead complex, abstract projects involving data analysis, modeling, and AI system evaluation.
- Design and execute evaluation protocols for machine learning models, including agreement metrics and failure-mode analysis.

AI/ML Research Engineer (Open Source Contributor)

OpenJarvis | 2026 - Present

- Contribute to a local-first personal AI agent framework (Python + Rust + Tauri), building on-device agents that run without cloud dependency.
- Design and refine prompts and system instructions for multi-agent orchestration workflows using LangGraph.
- Collaborate on AI safety and alignment: security guardrails, file access policy, capability-based RBAC, taint checking, and PII redaction for safe agent tool dispatch.
- Collaborate with an international open-source community (Stanford-affiliated project).

Graduate Researcher - Multimodal AI

UFPel | 2024 - Present

- Build a pipeline that detects and classifies 7 programming behaviors from student screen recordings using Vision-Language Models (Qwen2.5-VL via Ollama).
- Maintain a prompt library with versioned templates for VLM classification (structured output: label, confidence score, justification) across multiple model configurations (qwen, llama, gemma).
- Design evaluation rubrics and run agreement analysis against human annotation (Cohen's kappa, macro F1, confusion matrices), establishing annotation quality baselines for RLHF-ready training data.
- Implement gate sensitivity sweeps and bootstrap confidence intervals for statistical validation.

AI Agent Engineer

WMS Agent (Paipe) — Project | 2024 - Present

- Built a warehouse management AI agent using LangChain and LangGraph, integrating LLMs (Claude/GPT) with tool-calling for autonomous data querying.
- Designed evaluation framework: golden dataset (15 questions), out-of-scope refusal tests, and regression tests for prompt extraction attacks.
- Implemented LLMOps practices: structured logging (timestamp, source, provider, latency), audit trail, and monitoring for every agent interaction.
- Secured agent with rate limiting, input validation, CSV/formula injection prevention, and prompt injection defense.

ML Pipeline Engineer

Binary Defect Detection — Project | 2024

- Designed a 4-branch ML pipeline (classical texture, CNN fine-tuning, explainability, anomaly detection) for visual defect classification in manufacturing.
- Ran A/B comparison across architectures (ResNet50V2, EfficientNetB4, ConvNeXt-S) with Optuna hyperparameter search and MLflow experiment tracking.
- Built explainability reports using LIME and SHAP to analyze model failure modes.

Education

M.Sc. in Computer Science

Federal University of Pelotas (UFPel) | Apr 2024 - Present

B.Sc. in Computer Science

Federal University of Pelotas (UFPel) | Jun 2017 - Dec 2023

Technical Diploma in Electronics

Instituto Federal Sul-Rio-Grandense | Jun 2014 - Jun 2016

Skills

- **Generative AI:** LLM/VLM prompt engineering, system instructions, prompt library management, structured output, few-shot and chain-of-thought patterns
- **Agentic AI:** LangChain, LangGraph, multi-agent orchestration, tool calling, autonomous agents, local-first frameworks
- **LLMOps:** Experiment tracking (MLflow), structured logging, audit trails, monitoring, evaluation frameworks
- **Computer Vision:** OpenCV, VLMs (Qwen2.5-VL), image classification, defect detection, multimodal analysis
- **Machine Learning:** scikit-learn, PyTorch, TensorFlow, CNNs, anomaly detection
- **Programming:** Python (advanced), Rust (basic), SQL
- **Cloud:** AWS, Azure
- **Tools:** Docker, Git, Ollama, Jupyter
- **Languages:** Portuguese (native), English (advanced), Spanish (basic)